

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/763,076A

DATE: 06/29/2001

TIME: 12:10:06

Input Set : A:\PPD50348 US SEQ LIST.txt

Output Set: N:\CRF3\06292001\I763076A.raw

3 <110> APPLICANT: Broekaert, Willem
4 Francois, Isabelle
5 Evans, Ian
6 De Bolle, Miguel
7 Ray, John
9 <120> TITLE OF INVENTION: Genetic Method For The Expression of Polyproteins in Plants
11 <130> FILE REFERENCE: PPD50348/UST
C--> 13 <140> CURRENT APPLICATION NUMBER: US/09/763,076A
C--> 14 <141> CURRENT FILING DATE: 1999-08-17
16 <150> PRIOR APPLICATION NUMBER: GB 9818001.1
17 <151> PRIOR FILING DATE: 1998-08-18
19 <150> PRIOR APPLICATION NUMBER: GB 9826753.7
20 <151> PRIOR FILING DATE: 1998-12-14
22 <150> PRIOR APPLICATION NUMBER: PCT/GB99/02716
23 <151> PRIOR FILING DATE: 1999-08-17
25 <160> NUMBER OF SEQ ID NOS: 81
27 <170> SOFTWARE: PatentIn Ver. 2.1
29 <210> SEQ ID NO: 1
30 <211> LENGTH: 446
31 <212> TYPE: DNA
32 <213> ORGANISM: Dahlia merckii
34 <400> SEQUENCE: 1
35 atgggtgaatc ggtcgggttc gttctccgcg ttcgtttctga tccttttcgt gctcgccatc 60
36 tcagggttatac aaatcttttag ttcattttatt gaatatgata gtattttatat tctttttatgg 120
37 ttttatgtgt tctgacaagt tgcaaatatt gagtagatat cgcacccgtt agtggagaac 180
38 tatgcgagaa agctagcaag acatgggtcgg gaaactgttg caatacggga cattgtgaca 240
39 accaatgtaa atcatgggag ggtgcggccc atggagcgtg tcatgtgcgt aacgggaaac 300
40 acatgtgttt ctgttacttc aattgtaaaa aagccgaaaa gcttgctcaa gacaaactta 360
41 aagccgaaca actcgtcaa gacaaactta atgcccacaa gcttgaccgt gatgccaaga 420
42 aagtggttcc aaacgttgaa catccg 446
46 <210> SEQ ID NO: 2
47 <211> LENGTH: 118
48 <212> TYPE: PRT
49 <213> ORGANISM: Dahlia merckii
51 <400> SEQUENCE: 2
52 Met Val Asn Arg Ser Val Ala Phe Ser Ala Phe Val Leu Ile Leu Phe
53 1 5 10 15
55 Val Leu Ala Ile Ser Asp Ile Ala Ser Val Ser Gly Glu Leu Cys Glu
56 20 25 30
58 Lys Ala Ser Lys Thr Trp Ser Gly Asn Cys Gly Asn Thr Gly His Cys
59 35 40 45
61 Asp Asn Gln Cys Lys Ser Trp Glu Gly Ala Ala His Gly Ala Cys His
62 50 55 60
64 Val Arg Asn Gly Lys His Met Cys Phe Cys Tyr Phe Asn Cys Lys Lys
65 65 70 75 80
67 Ala Glu Lys Leu Ala Gln Asp Lys Leu Lys Ala Glu Gln Leu Ala Gln
68 85 90 95

ENTERED

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/763,076A

DATE: 06/29/2001

TIME: 12:10:06

Input Set : A:\PPD50348 US SEQ LIST.txt

Output Set : N:\CRF3\06292001\I763076A.raw

```

70 Asp Lys Leu Asn Ala Gln Lys Leu Asp Arg Asp Ala Lys Lys Val Val
71          100          105          110
73 Pro Asn Val Glu His Pro
74          115
78 <210> SEQ ID NO: 3
79 <211> LENGTH: 16
80 <212> TYPE: PRT
81 <213> ORGANISM: Artificial Sequence
83 <220> FEATURE:
84 <223> OTHER INFORMATION: Description of Artificial Sequence: Linker
85     propeptide
87 <400> SEQUENCE: 3
88 Ser Asn Ala Ala Asp Glu Val Ala Thr Pro Glu Asp Val Glu Pro Gly
89   1          5          10          15
93 <210> SEQ ID NO: 4
94 <211> LENGTH: 20
95 <212> TYPE: PRT
96 <213> ORGANISM: Artificial Sequence
98 <220> FEATURE:
99 <223> OTHER INFORMATION: Description of Artificial Sequence: Linker
100     propeptide
102 <400> SEQUENCE: 4
103 Lys Lys Ala Glu Lys Leu Ala Gln Asp Lys Leu Lys Ala Glu Gln Leu
104   1          5          10          15
106 Ile Gly Lys Arg
107          20
111 <210> SEQ ID NO: 5
112 <211> LENGTH: 40
113 <212> TYPE: PRT
114 <213> ORGANISM: Dahlia merckii
116 <400> SEQUENCE: 5
117 Lys Lys Ala Glu Lys Leu Ala Gln Asp Lys Leu Lys Ala Glu Gln Leu
118   1          5          10          15
120 Ala Gln Asp Lys Leu Asn Ala Gln Lys Leu Asp Arg Asp Ala Lys Lys
121          20          25          30
123 Val Val Pro Asn Val Glu His Pro
124          35          40
128 <210> SEQ ID NO: 6
129 <211> LENGTH: 44
130 <212> TYPE: PRT
131 <213> ORGANISM: Artificial Sequence
133 <220> FEATURE:
134 <223> OTHER INFORMATION: Description of Artificial Sequence: Linker
135     propeptide
137 <400> SEQUENCE: 6
138 Lys Lys Ala Glu Lys Leu Ala Gln Asp Lys Leu Lys Ala Glu Gln Leu
139   1          5          10          15
141 Ala Gln Asp Lys Leu Asn Ala Gln Lys Leu Asp Arg Asp Ala Lys Lys
142          20          25          30

```

RAW SEQUENCE LISTING

DATE: 06/29/2001

PATENT APPLICATION: US/09/763,076A

TIME: 12:10:06

Input Set : A:\PPD50348 US SEQ LIST.txt

Output Set: N:\CRF3\06292001\I763076A.raw

```

144 Val Val Pro Asn Val Glu His Pro Ile Gly Lys Arg
145      35      40
149 <210> SEQ ID NO: 7
150 <211> LENGTH: 20
151 <212> TYPE: PRT
152 <213> ORGANISM: Artificial Sequence
154 <220> FEATURE:
155 <223> OTHER INFORMATION: Description of Artificial Sequence: Linker
156     propeptide
158 <400> SEQUENCE: 7
159 Ala Ser Thr Thr Val Asp His Gln Ala Asp Val Ala Ala Thr Lys Thr
160   1      5      10      15
162 Ile Gly Lys Arg
163      20
167 <210> SEQ ID NO: 8
168 <211> LENGTH: 31
169 <212> TYPE: PRT
170 <213> ORGANISM: Amaranthus caudatus
173 <400> SEQUENCE: 8
174 Ala Ser Thr Thr Val Asp His Gln Ala Asp Val Ala Ala Thr Lys Thr
175   1      5      10      15
177 Ala Lys Asn Pro Thr Asp Ala Lys Leu Ala Gly Ala Gly Ser Pro
178      20      25      30
182 <210> SEQ ID NO: 9
183 <211> LENGTH: 522
184 <212> TYPE: DNA
185 <213> ORGANISM: Artificial Sequence
187 <220> FEATURE:
188 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
189     sequence
191 <220> FEATURE:
192 <221> NAME/KEY: CDS
193 <222> LOCATION: (76)..(513)
195 <400> SEQUENCE: 9
196 ctcgagtatt ttacaacaa ttaccaacaa caacaaacaa caaacaacat tacaattact 60
198 atttacaatt acacc atg gtg aat cgg tcg gtt gcg ttc tcc gcg ttc gtt 111
199           Met Val Asn Arg Ser Val Ala Phe Ser Ala Phe Val
200           1      5      10
202 ctg atc ctt ttc gtg ctc gcc atc tca gat atc gca tcc gtt agt gga 159
203 Leu Ile Leu Phe Val Leu Ala Ile Ser Asp Ile Ala Ser Val Ser Gly
204           15      20      25
206 gaa cta tgc gag aaa gct agc aag acg tgg tcg ggc aac tgt ggc aac 207
207 Glu Leu Cys Glu Lys Ala Ser Lys Thr Trp Ser Gly Asn Cys Gly Asn
208      30      35      40
210 acg gga cat tgt gac aac caa tgt aaa tca tgg gag ggt gcg gcc cat 255
211 Thr Gly His Cys Asp Asn Gln Cys Lys Ser Trp Glu Gly Ala Ala His
212  45      50      55      60
214 gga gcg tgt cat gtg cgt aac ggg aaa cac atg tgt ttc tgt tac ttc 303
215 Gly Ala Cys His Val Arg Asn Gly Lys His Met Cys Phe Cys Tyr Phe

```

RAW SEQUENCE LISTING

DATE: 06/29/2001

PATENT APPLICATION: US/09/763,076A

TIME: 12:10:06

Input Set : A:\PPD50348 US SEQ LIST.txt

Output Set: N:\CRF3\06292001\I763076A.raw

```

216          65          70          75
218 aat tgt tcc aac gct gct gac gag gtg gct acc cca gag gac gtg gag 351
219 Asn Cys Ser Asn Ala Ala Asp Glu Val Ala Thr Pro Glu Asp Val Glu
220          80          85          90
222 cca gga cag aag ttg tgc caa agg cca agt ggg aca tgg tca gga gtc 399
223 Pro Gly Gln Lys Leu Cys Gln Arg Pro Ser Gly Thr Trp Ser Gly Val
224          95          100          105
226 tgt gga aac aat aac gca tgc aag aat cag tgc att aga ctt gag aaa 447
227 Cys Gly Asn Asn Asn Ala Cys Lys Asn Gln Cys Ile Arg Leu Glu Lys
228          110          115          120
230 gca cga cat gga tct tgc aac tat gtc ttc cca gct cac aag tgt atc 495
231 Ala Arg His Gly Ser Cys Asn Tyr Val Phe Pro Ala His Lys Cys Ile
232 125          130          135          140
234 tgc tac ttt cct tgt taa taggagctc 522
235 Cys Tyr Phe Pro Cys
236          145
239 <210> SEQ ID NO: 10
240 <211> LENGTH: 145
241 <212> TYPE: PRT
242 <213> ORGANISM: Artificial Sequence
244 <220> FEATURE:
245 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
246 sequence
248 <400> SEQUENCE: 10
249 Met Val Asn Arg Ser Val Ala Phe Ser Ala Phe Val Leu Ile Leu Phe
250 1 5 10 15
252 Val Leu Ala Ile Ser Asp Ile Ala Ser Val Ser Gly Glu Leu Cys Glu
253 20 25 30
255 Lys Ala Ser Lys Thr Trp Ser Gly Asn Cys Gly Asn Thr Gly His Cys
256 35 40 45
258 Asp Asn Gln Cys Lys Ser Trp Glu Gly Ala Ala His Gly Ala Cys His
259 50 55 60
261 Val Arg Asn Gly Lys His Met Cys Phe Cys Tyr Phe Asn Cys Ser Asn
262 65 70 75 80
264 Ala Ala Asp Glu Val Ala Thr Pro Glu Asp Val Glu Pro Gly Gln Lys
265 85 90 95
267 Leu Cys Gln Arg Pro Ser Gly Thr Trp Ser Gly Val Cys Gly Asn Asn
268 100 105 110
270 Asn Ala Cys Lys Asn Gln Cys Ile Arg Leu Glu Lys Ala Arg His Gly
271 115 120 125
273 Ser Cys Asn Tyr Val Phe Pro Ala His Lys Cys Ile Cys Tyr Phe Pro
274 130 135 140
276 Cys
277 145
281 <210> SEQ ID NO: 11
282 <211> LENGTH: 534
283 <212> TYPE: DNA
284 <213> ORGANISM: Artificial Sequence
286 <220> FEATURE:

```

RAW SEQUENCE LISTING

DATE: 06/29/2001

PATENT APPLICATION: US/09/763,076A

TIME: 12:10:06

Input Set : A:\PPD50348 US SEQ LIST.txt

Output Set: N:\CRF3\06292001\I763076A.raw

```

287 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
288     sequence
290 <220> FEATURE:
291 <221> NAME/KEY: CDS
292 <222> LOCATION: (76)..(525)
294 <400> SEQUENCE: 11
295 ctcgagtatt tttaacaaca ttaccaacaa caacaacaa caacaacat tacaattact 60
297 attacaatt acacc atg gtg aat cgg tcg gtt gcg ttc tcc gcg ttc gtt 111
298     'Met Val Asn Arg Ser Val Ala Phe Ser Ala Phe Val
299         1             5             10
301 ctg atc ctt ttc gtg ctc gcc atc tca gat atc gca tcc gtt agt gga 159
302 Leu Ile Leu Phe Val Leu Ala Ile Ser Asp Ile Ala Ser Val Ser Gly
303         15             20             25
305 gaa cta tgc gag aaa gct agc aag acg tgg tcg ggc aac tgt ggc aac 207
306 Glu Leu Cys Glu Lys Ala Ser Lys Thr Trp Ser Gly Asn Cys Gly Asn
307         30             35             40
309 acg gga cat tgt gac aac caa tgt aaa tca tgg gag ggt gcg gcc cat 255
310 Thr Gly His Cys Asp Asn Gln Cys Lys Ser Trp Glu Gly Ala Ala His
311         45             50             55             60
313 gga gcg tgt cat gtg cgt aac ggg aaa cac atg tgt ttc tgt tac ttc 303
314 Gly Ala Cys His Val Arg Asn Gly Lys His Met Cys Phe Cys Tyr Phe
315         65             70             75
317 aat tgt aaa aaa gcc gaa aag ctt gct caa gac aaa ctt aaa gcc gaa 351
318 Asn Cys Lys Lys Ala Glu Lys Leu Ala Gln Asp Lys Leu Lys Ala Glu
319         80             85             90
321 caa ctc atc gga aag agg cag aag ttg tgc caa agg cca agt ggg aca 399
322 Gln Leu Ile Gly Lys Arg Gln Lys Leu Cys Gln Arg Pro Ser Gly Thr
323         95             100             105
325 tgg tca gga gtc tgt gga aac aat aac gca tgc aag aat cag tgc att 447
326 Trp Ser Gly Val Cys Gly Asn Asn Asn Ala Cys Lys Asn Gln Cys Ile
327         110             115             120
329 aga ctt gag aaa gca cga cat gga tct tgc aac tat gtc ttc cca gct 495
330 Arg Leu Glu Lys Ala Arg His Gly Ser Cys Asn Tyr Val Phe Pro Ala
331         125             130             135             140
333 cac aag tgt atc tgc tac ttt cct tgt taa taggagctc 534
334 His Lys Cys Ile Cys Tyr Phe Pro Cys
335         145
338 <210> SEQ ID NO: 12
339 <211> LENGTH: 149
340 <212> TYPE: PRT
341 <213> ORGANISM: Artificial Sequence
343 <220> FEATURE:
344 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
345     sequence
347 <400> SEQUENCE: 12
348 Met Val Asn Arg Ser Val Ala Phe Ser Ala Phe Val Leu Ile Leu Phe
349     1             5             10             15
351 Val Leu Ala Ile Ser Asp Ile Ala Ser Val Ser Gly Glu Leu Cys Glu
352         20             25             30

```

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

VERIFICATION SUMMARY

DATE: 06/29/2001

PATENT APPLICATION: US/09/763,076A

TIME: 12:10:07

Input Set : A:\PPD50348 US SEQ LIST.txt

Output Set: N:\CRF3\06292001\I763076A.raw

L:13 M:270 C: Current Application Number differs, Replaced Application Number

L:14 M:271 C: Current Filing Date differs, Replaced Current Filing Date

L:414 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:13

L:2059 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:52

L:2251 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:66